

Integrar a rede de repositórios RCAAP no ecossistema de gestão de ciência – PTCRIS

Paulo Lopes, FCT, plopes@fccn.pt

José Carvalho, Universidade do Minho, jcarvalho@sdum.uminho.pt

Fernando Ribeiro, FCT, fernando.ribeiro@fccn.pt

Paulo Graça, FCT, paulo.graca@fccn.pt

Resumo da Proposta

Os repositórios institucionais são parte importante do ecossistema de gestão de ciência. Para integrar estes sistemas num ecossistema mais amplo, é necessário garantir o cumprimento de um quadro normativo que garanta a interoperabilidade entre os sistemas.

Esta comunicação tem como principal objetivo dar a conhecer o trabalho desenvolvido ao nível dos repositórios digitais no sentido de aumentar o seu nível de interoperabilidade de forma a facilitar os processos de workflow dos utilizadores nas diversas plataformas do ecossistema.

A comunicação abordará as cinco possíveis interações identificadas; a) A possibilidade de um autor reclamar a sua identidade ou os trabalhos depositados num Repositório; b) A possibilidade de um autor efetuar um depósito num repositório a partir de um sistema de gestão de ciência (CRIS) como, por exemplo, o sistema nacional de gestão curricular; c) a adoção de mecanismos de sincronização da informação; d) a adoção de mecanismos de controlo de autoridade para identificar univocamente autores, organizações e financiamento; e, e) a definição de tarefas de curadoria com o objetivo de atualizar a informação existente nos repositórios de forma a dotá-la de identificadores persistentes de autores, organizações, financiamento ou de vocabulários e semânticas comuns aos restantes sistemas do ecossistema de gestão e ciência.

Serão apresentados exemplos práticos do trabalho efetuado e abordadas as questões estratégicas relacionadas com o esse trabalho (desafios, necessidades, cenários e metodologias).

Tipo de Trabalho

- Comunicação

Tema da Conferência

- Gestão de informação de Ciência e Tecnologia

- CRIS – Sistemas de Gestão de informação de Ciência e Tecnologia
- Interoperabilidade entre sistemas de informação de apoio à atividade científica e académica
- Normas e diretrizes
- Identificadores persistentes

Palavras-chave

Interoperabilidade; normas e diretrizes; repositórios; sistema de curricula

Audiência

Gestores de repositórios, bibliotecários, gestores de dados de investigação, programadores, gestores de ciência, gestores de sistemas de gestão de ciência

Proposta

A gestão de ciência e a tecnologia e ensino superior tem vindo, ao longo do tempo, a tornar-se cada vez mais complexa. Contribuem para esta complexidade, entre outros fatores, o aumento do número de agentes, instrumentos e a quantidade de informação envolvidos nos ciclos do processo de investigação e dos processos do ensino superior.

Para fazer face a esta complexidade e ao seu aumento, têm vindo a ser desenvolvidos e adotados, em todo o mundo, sistemas integrados de ciência, tecnologia e de ensino.

Em Portugal, as Instituições de Ensino Superior (IES) têm feito esforços notáveis na adoção destes sistemas. Contudo, estes sistemas, com diferentes níveis de maturidade, tendem a operar de forma isolada. Neste contexto, em 2014 surge o programa PTCRIS¹ (Portuguese Current Research Information System). O PTCRIS é um programa desenvolvido pela Unidade de Computação Científica Nacional (UCCN) da FCT, que tem como objetivo o desenvolvimento sustentado de um ecossistema integrado de informação de suporte à atividade científica nacional. O desenvolvimento de tal ecossistema é crucial para garantir, por um lado, a desburocratização e a simplificação dos processos administrativos, por outro, a produção, gestão e acesso a dados fidedignos, completos e atualizados, dos quais depende a efetividade e eficiência dos processos de tomada de decisão.

Para dar cumprimento ao seu objetivo, o PTCRIS dedica-se à definição de um quadro normativo e à criação de infraestruturas que viabilizem a adoção desse mesmo normativo (Figura 1):

¹ <https://ptcris.pt/>



Figura 1 – Âmbito de atuação do PTCRIS

O RCAAP² é o principal instrumento para a implementação do Open Access (OA) em Portugal. Tem por missão promover, apoiar e facilitar a adoção do acesso aberto ao conhecimento científico em Portugal e tem como objetivo a recolha, agregação e indexação dos conteúdos científicos em acesso aberto (ou acesso livre) existentes nos repositórios institucionais das entidades nacionais de ensino superior e outras organizações de I&D.

Como serviço nacional que agrega a produção científica de vários repositórios institucionais e revistas científicas, o RCAAP constitui-se como peça importante do ecossistema PTCRIS (ver Figura 2). A adoção pelo RCAAP dos normativos de interoperabilidade definidos pelo PTCRIS são cruciais para a sua integração com outros sistemas de gestão de ciência.

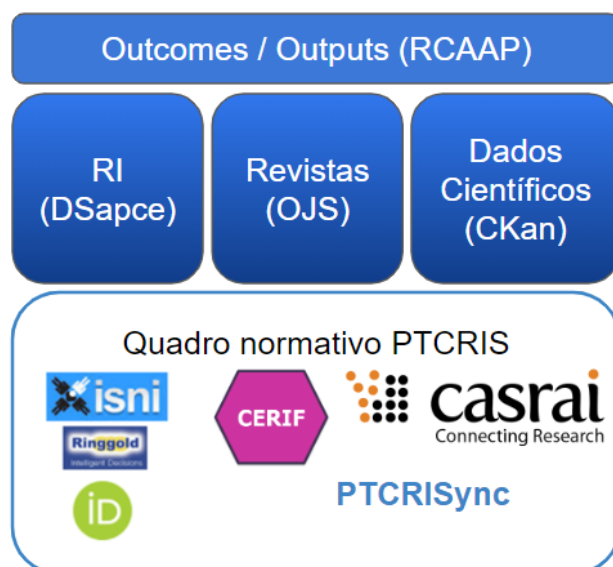


Figura 2 - Adopção no RCAAP do quadro normativo do PTCRIS

² <http://projeto.rcaap.pt/>

Tendo sempre como ponto de referência o utilizador (autores, investigadores, instituições), o aumento da interoperabilidade na rede RCAAP alinha-se com a visão geral do PTCRIS que defende o princípio “*enter once, re-use multiple*”.

De forma a promover este princípio no RCAAP, foram identificados cinco blocos de casos de uso (Figura 3) relacionados com necessidades de interação entre os utilizadores e sistemas ou apenas entre sistemas (trocas de informação máquina - máquina):

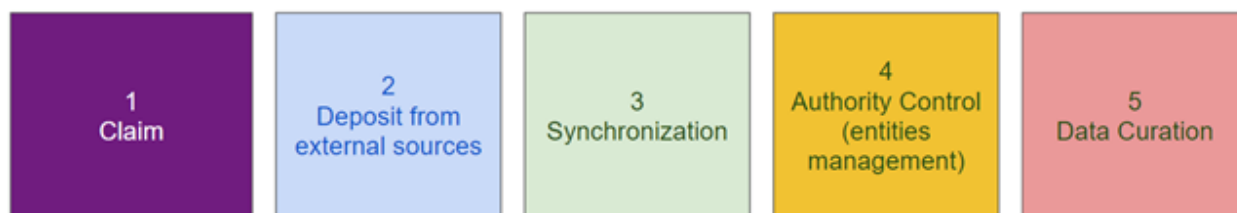


Figura 3 – Blocos de casos de uso de interoperabilidade

Cada um destes blocos prevê um conjunto de ações que promovem a interoperabilidade tendo por base o quadro normativo definido pelo PTCRIS:

Bloco 1 – Claim – Contempla a possibilidade de um autor: a) relacionar a sua conta Ciência ID (o Ciência ID é um identificador único para investigadores em Portugal) com a sua conta local no seu repositório institucional; e, b) a possibilidade de reclamar os seus trabalhos no repositório. A Figura 4 contextualiza o primeiro caso e mostra como a possibilidade de login via Ciência ID foi acrescentada ao repositório. O processo de login por esta via envolve o relacionamento da conta local com a conta Ciência ID. Caso esse relacionamento não seja validado é criada uma nova conta no repositório:

Figura 4 – Login no repositório com Ciência ID

A Figura 5 contextualiza o segundo caso. Do lado do sistema de gestão curricular (Ciência Vitae³) foi criado um método de importação de trabalhos a partir do RCAAP. O utilizador pode seleccionar vários critérios para isolar a sua produção científica - busca por autor, por título de trabalho, por repositório, por tipo de produção, por datas, entre outros. A API do portal RCAAP está a ser desenvolvida para permitir também a importação de trabalhos por identificador de autor (Ciência ID e/ou ORCID).

³ <https://www.cienciavitae.pt/>

Importador RCAAP

Critério
 Repositório ▼

Termo a pesquisar
 Universidade Atlântica (Repositório Científico A ▼) (+)

PESQUISAR

Figura 5 - Importação de trabalhos a partir de plataforma Ciência Vitae

Bloco 2 – Depósito no Repositório a partir de entidades externas – Contempla a possibilidade de se enviar uma publicação para o Repositório a partir de um sistema externo como, por exemplo, o sistema de gestão de currículos (Ciência Vitae). A Figura 6 mostra o último passo de um processo de envio de uma publicação para um repositório a partir da Plataforma Ciência Vitae:

CIÊNCIAVITAE Development Vitae Bruno Monteiro

IDENTIFICAÇÃO FORMAÇÃO PERCURSO PROFISSIONAL PROJETOS **PRODUÇÕES** ATIVIDADES DISTINÇÕES

Depositar Produção

Está a depositar *Maximizing Project Value: A Project Manager's Guide*

Resumo dos dados do depósito

Verifique, por favor, se todas as opções do depósito estão em conformidade com o pretendido. Pode alterar uma ou mais opções clicando no ícone correspondente. O processo de depósito termina somente quando clicar na opção "concluir depósito".

Repositório para depósito	Universidade de Lisboa - Instituto Superior Técnico	
Colecção	Genérica	
Tipo de acesso ao documento	Acesso Embargado até 2020.04.08	
Ficheiro a anexar	mpv_pmg_final.pdf	
Informação de financiamento	821300 - Hairmate - Hybrid Aircraft seating Manuf... FCT/PPCT/00 - Hybrid Aircraft seating Manufacturing...	

ANTERIOR CANCELAR CONCLUIR DEPÓSITO

Figura 6 - Depósito a partir da plataforma Ciência Vitae

Bloco 3 – Sincronização – Prevê a adoção do mecanismo de sincronização PTCRISync⁴ nos repositórios e no Portal RCAAP. Este mecanismo já funciona em alguns sistemas do ecossistema, como é o caso da Plataforma Ciência Vitae. Na figura 7 podemos ver o ecrã onde são identificadas as produções já sincronizadas ou por sincronizar com o ORCID. Neste caso o ORCID funciona com um *Hub* central que permite fazer circular a informação entre os diversos sistemas que adotem este método de sincronização. Ao referenciar a sua produção científica (no Repositório, no seu perfil ORCID ou no seu CV) o utilizador pode fazer circular essa informação entre os vários sistemas sem ter a necessidade de reintroduzir a informação em cada um dos sistemas.

Sincronizar com ORCID ?

Data mais recente Ver todos os tipos 20

Por sincronizar (0) Alertas (0) Sincronizadas (5)

☐ Selecionar tudo Última atualização

☐ Parar sincronização Curadoria e integração de sistemas 12 abril às 16:42h

Artigo em conferência apresentado na conferência realizada em 2018
 source-work-id: cv-prod-id-274820; handle: http://hdl.handle.net/1822/56590

Figura 7 - Sincronização de informação com o perfil ORCID

Bloco 4 – Controlo de autoridade – Prevê a introdução de mecanismos de controlo de autoridade para autores, instituições e financiamento. A figura 8 mostra como nos repositórios foi adicionada a possibilidade de relacionar um autor com o os seus identificadores (Ciência ID e/ou ORCID) no ato do depósito de uma publicação:

Pesquisa na autoridade de nomes

Procurar por identificador

CIENCIA ID/ORCID

Lista de Resultados

Laranjeira, Cátia
 Laranjeira (Teste), Cátia
 'Laranjeira, Cátia' (não em Autoridade)

Apresentar mais resultados

Aceitar Cancelar

Informação sobre o autor

Nome
 Laranjeira, Cátia

Variantes

Afiliação

Identificadores
 3415-1354-5489 CIENCIA ID
 0000-0003-3952-8294 ORCID
 23091392500 Scopus

Documentos do autor no Repositório

Figura 8 - Interface para seleção de autor (com identificador)

⁴ <https://ptcris.pt/ptcrisync/>

Bloco 5 – Tarefas de curadoria – Define tarefas de curadoria a executar com o objetivo de enriquecer as publicações existentes nos repositórios de forma a dotá-las com informação sobre identificação de autores, instituições e financiamento.

Partindo das linhas de ação definidas e tendo por base os blocos atrás descritos, efetuou-se o seguinte trabalho:

Repositórios digitais – A intervenção fez-se ao nível dos Repositórios Institucionais alojados no serviço SARI (serviço de alojamento de repositórios do RCAAP) que tem por base a versão 5 do DSpace. Foram acrescentadas as seguintes capacidades:

- Introdução da possibilidade de autenticação através da conta Ciência ID;
- Criação de uma interface que permite, no acto do depósito, associar o nome de um autor com o seu identificador (Ciência ID ou ORCID);
- Alteração dos formulários de depósito de forma a ficarem conformes com as novas Guidelines 4 do OpenAIRE e com as definições COAR no que respeita a tipologias de trabalhos, tipos de acesso e versões dos documentos;
- Alteração da interface OAI-PMH de forma a expor a informação nos novos formatos;
- Criados novos fluxos de informação para suportar o depósito a partir de entidades externas.

Paralelamente a equipa do RCAAP tem vindo a trabalhar conjuntamente com a Duraspace de forma a dotar a versão 7 do DSpace com capacidade de gerir novas entidades (autores, organizações e financiamento) e de suportar nativamente o esquema de metadados de acordo com as guidelines 4 do OpenAIRE.

Portal RCAAP - O Portal RCAAP foi renovado em 2018, tendo por base o software do LaReferencia, com o objetivo principal de acabar com a limitação de suporte de esquemas de metadados hierárquicos que se verificava na versão anterior.

O suporte de novos formatos prevê que o Portal seja capaz de suportar a agregação, gestão e exposição de entidades CRIS: Autores, Organizações e Financiamento. Para isso está a ser efetuado um trabalho conjunto com a equipa do LaReferencia de forma a alterar o modelo de dados do software para que possa gerir as novas entidades. Está também a ser desenvolvida uma nova versão da API para exposição deste tipo de informação.

A utilização de identificadores comuns para estas entidades vai permitir, por exemplo, que o portal possa apresentar páginas de perfis de autores tendo por base a informação agregada a partir dos repositórios (publicações, organizações, financiamento) podendo essas páginas serem ainda enriquecidas com informação adicional proveniente do Curriculum Vitae desse autor presente no sistema nacional de curricula.

Trabalho futuro

O plano de trabalhos para o ano 2019 prevê duas atividades principais:

- Continuar a implementar os casos de uso definidos nos cinco blocos de integrações RCAAP – PTCRIS;
- Apoiar os desenvolvimentos do DSpace7. Como referido, o trabalho efetuado iniciou-se na versão 5 do DSpace mas prevê a migração para a nova versão assim que esta estiver disponível.

CONCLUSÕES

O objetivo de integrar a rede RCAAP no ecossistema de gestão de ciência PTCRIS tem esbarrado em inúmeras dificuldades, nomeadamente com o facto de as funcionalidades dos softwares de gestão de repositórios e revistas estarem concentradas na gestão de publicações, sem estarem preparados para gerir outras entidades importantes do ecossistema de gestão de ciência como é o caso dos autores, das organizações ou do financiamento.

De forma a ultrapassar estas dificuldades foram adotadas duas grandes linhas de ação:

- Por um lado, trabalhar conjuntamente com a comunidade responsável pelos principais softwares utilizados pelo RCAAP de forma a influenciar o seu ciclo de desenvolvimento para que os mesmos integrem funcionalidades de gestão de novas entidades e de adotem normas internacionais de classificação dos dados;
- Por outro, modificar e adaptar os softwares, ferramentas e processo em utilização de forma a que possam responder às necessidades de interoperabilidade que se pretendem.

Pretende-se que este trabalho represente a visão de uma rede organizada de repositórios institucionais que combina um alinhamento nacional e internacional em relação às diretrizes e organização das informações dos diferentes sistemas. Os desenvolvimentos focam-se na reutilização da informação, não apenas para o gestor de repositórios, mas também, e sobretudo, para o próprio utilizador, integrando e permitindo a sincronização de informações de diferentes fontes. Por fim, introduz o conceito de gestão de novas entidades, para além das publicações científicas, através dos diferentes sistemas.

Espera-se que a adoção de normativos nacionais e internacionais comuns aos diversos sistemas de gestão de ciência e a gestão destas novas entidades (autores, organizações e financiamento) pelos repositórios tragam benefícios a diversos níveis:

Para os autores - Fácil acesso a informação relevante sobre oportunidades de financiamento, pessoas, parceiros e/ou concorrentes. Permite identificar projetos, publicações e infraestruturas científicas. Gestão simplificada da sua informação através da introdução da informação uma vez e reutilização em múltiplas em sistemas do ecossistema PTCRIS. Agilização da compilação de dados para relatórios a partir de informação previamente introduzida no ecossistema.

Para as organizações - Facilitação da atividade de *reporting* uma vez que a integração de sistemas facilita a reunião de informação autoritativa, completa e fidedigna diversa sobre a prática científica tornando mais célere a produção de relatórios. Facilitação do acesso a indicadores/medidas do desempenho permitindo uma avaliação do desempenho de investigadores /instituições mais fácil e rápida. Otimização dos processos de financiamento uma vez que a existência de informação autoritativa, completa e fidedigna e atualizada sobre a atividade de investigação permite otimizar, quer do ponto de vista de concessão, quer de obtenção, o processo de financiamento.

Os trabalhos devem ser enviados no formato atual através do sistema de submissões em:

<https://conferencias.rcaap.pt/confoa/>